Predictive Analytics Exam—October 2025

The Predictive Analytics exam is administered as a three hour and thirty-minute exam requiring analysis of data in the context of a business problem and submission of written responses to specified tasks. There is no scheduled break for candidates. Candidates will have access to a computer equipped with Microsoft Word and Microsoft Excel.[1] **Note that starting with the April 2023 administration, R and RStudio will NOT be available**. Written responses will be submitted electronically via Microsoft Word. For additional details, please refer to the Exam PA home page.

Exam PA assumes knowledge of probability, mathematical statistics, and selected analytical techniques as covered in Exam P (Probability), VEE Mathematical Statistics, and Exam SRM (Statistics for Risk Modeling).

Please check the Updates section on this exam's home page for any changes to the exam or syllabus.

The learning objectives and outcomes provided on the following pages follow the five modules of the e-Learning support provided. There is an additional module that is an Introduction to R. That knowledge is not part of the learning objectives for the course, but is needed to be able to work with the material in the other modules. The ranges of weights shown in the Learning Objectives below are intended to apply to the large majority of exams administered. On occasion, the weights of topics on an individual exam may fall outside the published range. Candidates should also recognize that tasks often cover multiple learning objectives, including some weight for communication in most tasks.

---

[1] The Prometric computers will have the 2016 versions of Microsoft Word and Excel.

LEARNING OBJECTIVES

1. **Predictive Analytics Problem Definition (10-20%)**
Learning Objectives
The Candidate will be able to identify the business problem, how the available data relates to possible analyses, and use the information to propose an appropriate modeling approach.
Learning Outcomes
The Candidate will be able to: a) Assess whether descriptive, predictive, and prescriptive analytics applies to a business problem. b) Describe the characteristics of predictive modeling problems. c) Explain the concepts of bias, variance, model complexity, and the bias-variance trade-off. d) Translate a vague question into one that can be analyzed with statistics and predictive analytics to solve a business problem. e) Consider factors such as available data and technology, significance of business impact, and implementation challenges to define the problem. f) Assess what additional information or next steps would improve the ability to apply predictive analytics to a business problem.

2. **Topic: Data Exploration and Visualization (20-30%)**
Learning Objectives
The Candidate will be able to work with various data types, understand principles of data design, and construct a variety of common visualizations for exploring data.
Learning Outcomes
The Candidate will be able to: a) Identify structured and unstructured data types. b) Identify the types of variables and terminology used in predictive modeling. c) Evaluate effective data design with respect to time frame, sampling, and granularity. d) Apply the key principles of constructing graphs. e) Apply univariate data exploration techniques. f) Apply bivariate data exploration techniques.

3. **Topic: Data Transformations and Unsupervised Learning Techniques (10-20%)**
Learning Objectives
The Candidate will be able to transform data, including application of unsupervised learning techniques, to improve predictive modeling outcomes.
Learning Outcomes

The Candidate will be able to:
 	
   a) Create features from existing data that may add value.
 	
   b) Apply principal components analysis to transform data.
 	
   c) Apply K-means and hierarchical clustering to transform data.

4. **Topic: Generalized Linear Models (25-35%)**
Learning Objectives
The Candidate will be able to select and validate a generalized linear model (GLM) and apply related modeling concepts as appropriate for a given business problem.
Learning Outcomes

The Candidate will be able to:
 	
   a) Select and validate a GLM as appropriate for a business problem.
 	
   b) Apply offsets and weights as appropriate.
 	
   c) Interpret model coefficients, including interaction terms.
 	
   d) Select appropriate hyperparameters for regularized regression.

5. **Topic: Tree-Based Models (10-20%)**
Learning Objectives
The Candidate will be able to apply tree-based models as appropriate for a business problem.
Learning Outcomes

The Candidate will be able to:
 	
    a) Construct, prune, and validate regression and classification trees.
 	
    b) Apply bagging and random forests as appropriate.
 	
    c) Apply boosting as appropriate.

REQUIRED RESOURCES:

e-Learning Modules

Candidates will have access to a series of five-Learning modules providing instruction in the objectives stated above. The modules will also provide guidance regarding knowledge and approaches that will be expected in the exam.

R and RStudio

To complete the modules, candidates will be expected to be able to work with R within the RStudio environment. For those unfamiliar with the environment, instruction is provided in the first e-Learning module. The code in the modules was tested on version 4.2 of R using package versions from August 2022. Versions 4.0 and later should also work. Because R and RStudio will not be available at the examination, all code and output relevant to the tasks will be provided as part of the exam materials. Candidates are encouraged to work through all the R examples in the modules, understanding the code and output. Facility in R will also be useful as preparation for the Advanced Topics in Predictive Analytics assessment.

Textbooks

There are four texts required for the course. Two of them are the texts for the Statistics for Risk Modeling Exam. These are listed below this paragraph. It is assumed that candidates are familiar with all this material. Explicit reference to parts of these texts may be made from time to time within the modules. However, that does not imply the other sections are unimportant.

Regression Modeling with Actuarial and Financial Applications, Edward W. Frees, 2010, New York: Cambridge. ISBN: 978-0-521-13596-2. Note that while Chapters 7-9 are covered in Exam SRM, these models will not be covered in this exam.

Chapter 1 – Background only
Chapter 2 – Sections 1-8
Chapter 3 – Sections 1-5
Chapter 5 – Sections 1-7
Chapter 6 – Sections 1-3
Chapter 11 – Sections 1-6
Chapter 12 – Sections 1-4
Chapter 13 – Sections 1-6

An Introduction to Statistical Learning, with Applications in R, James, Witten, Hastie, Tibshirani, 2013, New York: Springer. There is now a second edition, 2021. PDFs of both editions can be obtained via links found here: https://www.statlearning.com. The material used on this exam is virtually unchanged and so either version may be used. The following table indicates the relevant sections from the two editions.

First Edition	Second Edition
Chapter 2 – Sections 1-3	Chapter 2 – Sections 1-3
Chapter 3 – Sections 1-6	Chapter 3 – Sections 1-6
Chapter 4 – Sections 1-3 and parts of 6 (referenced in Module 6, not on SRM)	Chapter 4 – Sections 1-3 and parts of 6 (referenced in Module 6, not on SRM)
Chapter 5 – Sections 1 and 3 (excluding 5.3.4)	Chapter 5 – Sections 1 and 3 (excluding 5.3.4)
Chapter 6 – Sections 1-7	Chapter 6 – Sections 1-5
Chapter 8 – Sections 1-3	Chapter 8 – Sections 1-3 (excluding 8.2.4, 8.2.5, and 8.3.5
Chapter 10 – Sections 1-6	Chapter 12 – Sections 1-2, 4-5 (excluding 12.5.2)

The two additional texts required for this exam are shown below. The indicated chapters are referenced in the modules, but not all of each chapter is required. The modules will provide further guidance.

R for Everyone, 2nd ed. Lander, 2017, Boston: Addison-Wesley, ISBN 978-0-13-454692-6.

Chapters 1-10, 14, 26 and 28

Data Visualization: A Practical Introduction, Healy, 2018, Princeton University Press, ISBN 978-0-691-18162-2. This book may be available as web pages at http://socviz.co/. Note that this version can be viewed only on the web, there is no PDF version to download. The author has indicated that at some point the web version may be discontinued

Chapters 1-4

Past Exams and Model Solutions

Past exams and model solutions can be found here. Note that older exams do not use the same format or learning objectives as this exam.

A Note About Shortcut Keys

Due to security issues, many shortcut key combinations will be disabled on the Prometric computers. Click here to obtain a list of disabled shortcut keys.

Problem Statements

A hardcopy of the problem statements will not be available at Prometric testing centers. The statements will be available for the entirety of the exam on-screen.